

SEQUENCE LISTING

<110> Heidner, Hans Walter
 Klimstra, William Brown
 Ryman, Katherine Diana

<120> METHODS AND COMPOSITIONS COMPRISING PROTEIN L IMMUNOGLOBULIN
 BINDING DOMAINS FOR CELL-SPECIFIC TARGETING

<130> 9568-2

<140> US 10/593,841
 <141> 2007-05-17

<150> PCT/US2004/013281
 <151> 2004-04-29

<160> 27

<170> PatentIn version 3.3

<210> 1
 <211> 2657
 <212> DNA
 <213> Peptostreptococcus magnus

<220>
 <221> CDS
 <222> (301)..(2460)
 <223> protein L

<220>
 <221> repeat_unit
 <222> (592)..(819)
 <223> PpL1

<220>
 <221> repeat_unit
 <222> (820)..(1035)
 <223> PpL2

<220>
 <221> repeat_unit
 <222> (1036)..(1251)
 <223> PpL3

<220>
 <221> repeat_unit
 <222> (1252)..(1467)
 <223> PpL4

<220>
 <221> repeat_unit
 <222> (1468)..(1686)
 <223> PpL5

```

<400> 1
atagtgtctac aataaaggat ggcactgaac taccgaacct aacgcttgtg gacattgtct 60

tgggaaattt ggacagtgga cgaatcaaga acaccattaa ttaaattggt gaagttcgat 120

tgttgatcac cctttttggg taaacaataa ctaaggaatg gcaatatatt tgcttgaaa 180

cgaatttgat ttaaatagca ttgaatgcaa aaaaatttaa aaggaggaga caaattccac 240

ccttattaga agggaagttt ccattgtcat gatattatga aaattaataa gaaattatta 300

atg gct gca ctt gca ggt gca att gta gta aca ggt gga gta gga tct 348
Met Ala Ala Leu Ala Gly Ala Ile Val Val Thr Gly Gly Val Gly Ser
1 5 10 15

tac gca gct gat gaa cct att gat ctt gaa aaa ctt gaa gaa aaa agg 396
Tyr Ala Ala Asp Glu Pro Ile Asp Leu Glu Lys Leu Glu Glu Lys Arg
20 25 30

gat aaa gaa aat gta gga aat tta cca aaa ttc gat aat gaa gtt aaa 444
Asp Lys Glu Asn Val Gly Asn Leu Pro Lys Phe Asp Asn Glu Val Lys
35 40 45

gat ggt tca gaa aat cca atg gct aaa tat cca gat ttc gat gat gaa 492
Asp Gly Ser Glu Asn Pro Met Ala Lys Tyr Pro Asp Phe Asp Asp Glu
50 55 60

gcc agt aca aga ttt gaa aca gaa aac aat gaa ttt gaa gaa aaa aaa 540
Ala Ser Thr Arg Phe Glu Thr Glu Asn Asn Glu Phe Glu Glu Lys Lys
65 70 75 80

gtt gtt tct gat aac ttt ttt gat caa tca gaa cat ccg ttt gta gaa 588
Val Val Ser Asp Asn Phe Phe Asp Gln Ser Glu His Pro Phe Val Glu
85 90 95

aat aaa gaa gaa aca cca gaa aca cca gaa act gat tca gaa gaa gaa 636
Asn Lys Glu Glu Thr Pro Glu Thr Pro Glu Thr Asp Ser Glu Glu Glu
100 105 110

gta aca atc aaa gct aac cta atc ttt gca aat gga agc aca caa act 684
Val Thr Ile Lys Ala Asn Leu Ile Phe Ala Asn Gly Ser Thr Gln Thr
115 120 125

gca gaa ttc aaa gga aca ttt gaa aaa gca aca tca gaa gct tat gcg 732
Ala Glu Phe Lys Gly Thr Phe Glu Lys Ala Thr Ser Glu Ala Tyr Ala
130 135 140

tat gca gat act ttg aag aaa gac aat gga gaa tat act gta gat gtt 780
Tyr Ala Asp Thr Leu Lys Lys Asp Asn Gly Glu Tyr Thr Val Asp Val
145 150 155 160

gca gat aaa ggt tat act tta aat att aaa ttt gct gga aaa gaa aaa 828
Ala Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly Lys Glu Lys
165 170 175

aca cca gaa gaa cca aaa gaa gaa gtt act att aaa gca aac tta atc 876
Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala Asn Leu Ile

```

180								185				190				
tat	gca	gat	gga	aaa	aca	caa	aca	gca	gaa	ttc	aaa	gga	aca	ttt	gaa	924
Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	Thr	Phe	Glu	
		195						200				205				
gaa	gca	aca	gca	gaa	gca	tac	aga	tat	gca	gat	gca	tta	aag	aag	gac	972
Glu	Ala	Thr	Ala	Glu	Ala	Tyr	Arg	Tyr	Ala	Asp	Ala	Leu	Lys	Lys	Asp	
	210					215				220						
aat	gga	gaa	tat	aca	gta	gac	gtt	gca	gat	aaa	ggt	tat	act	tta	aat	1020
Asn	Gly	Glu	Tyr	Thr	Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr	Thr	Leu	Asn	
225					230					235					240	
att	aaa	ttt	gct	gga	aaa	gaa	aaa	aca	cca	gaa	gaa	cca	aaa	gaa	gaa	1068
Ile	Lys	Phe	Ala	Gly	Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	
			245						250					255		
gtt	act	att	aaa	gca	aac	tta	atc	tat	gca	gat	gga	aaa	aca	caa	aca	1116
Val	Thr	Ile	Lys	Ala	Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	
			260					265					270			
gca	gaa	ttc	aaa	gga	aca	ttt	gaa	gaa	gca	aca	gca	gaa	gca	tac	aga	1164
Ala	Glu	Phe	Lys	Gly	Thr	Phe	Glu	Glu	Ala	Thr	Ala	Glu	Ala	Tyr	Arg	
		275					280					285				
tat	gct	gac	tta	tta	gca	aaa	gaa	aat	ggt	aaa	tat	aca	gta	gac	gtt	1212
Tyr	Ala	Asp	Leu	Leu	Ala	Lys	Glu	Asn	Gly	Lys	Tyr	Thr	Val	Asp	Val	
	290					295				300						
gca	gat	aaa	ggt	tat	act	tta	aat	att	aaa	ttt	gct	gga	aaa	gaa	aaa	1260
Ala	Asp	Lys	Gly	Tyr	Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly	Lys	Glu	Lys	
305					310					315					320	
aca	cca	gaa	gaa	cca	aaa	gaa	gaa	gtt	act	att	aaa	gca	aac	tta	atc	1308
Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala	Asn	Leu	Ile	
				325					330					335		
tat	gca	gat	gga	aaa	act	caa	aca	gca	gag	ttc	aaa	gga	aca	ttt	gca	1356
Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	Thr	Phe	Ala	
			340					345					350			
gaa	gca	aca	gca	gaa	gca	tac	aga	tac	gct	gac	tta	tta	gca	aaa	gaa	1404
Glu	Ala	Thr	Ala	Glu	Ala	Tyr	Arg	Tyr	Ala	Asp	Leu	Leu	Ala	Lys	Glu	
		355					360					365				
aat	ggt	aaa	tat	aca	gca	gac	tta	gaa	gat	ggt	gga	tac	act	att	aat	1452
Asn	Gly	Lys	Tyr	Thr	Ala	Asp	Leu	Glu	Asp	Gly	Gly	Tyr	Thr	Ile	Asn	
	370					375				380						
att	aga	ttt	gca	ggt	aag	aaa	gtt	gac	gaa	aaa	cca	gaa	gaa	aaa	gaa	1500
Ile	Arg	Phe	Ala	Gly	Lys	Lys	Val	Asp	Glu	Lys	Pro	Glu	Glu	Lys	Glu	
385					390					395					400	
caa	gta	aca	att	aaa	gaa	aat	ata	tat	ttt	gaa	gat	gga	aca	gta	caa	1548
Gln	Val	Thr	Ile	Lys	Glu	Asn	Ile	Tyr	Phe	Glu	Asp	Gly	Thr	Val	Gln	
				405					410					415		

act gca aca ttt aaa gga aca ttt gca gaa gcg aca gca gaa gca tac	1596
Thr Ala Thr Phe Lys Gly Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr	
420 425 430	
aga tat gca gat ttg tta tca aaa gaa cat ggt aaa tac aca gca gac	1644
Arg Tyr Ala Asp Leu Leu Ser Lys Glu His Gly Lys Tyr Thr Ala Asp	
435 440 445	
ttg gaa gat ggt gga tac act atc aac att aga ttt gct gga aaa gaa	1692
Leu Glu Asp Gly Gly Tyr Thr Ile Asn Ile Arg Phe Ala Gly Lys Glu	
450 455 460	
gaa cca gaa gaa aca cca gaa aaa cca gaa gta caa gac gga tat gca	1740
Glu Pro Glu Glu Thr Pro Glu Lys Pro Glu Val Gln Asp Gly Tyr Ala	
465 470 475 480	
tca tac gaa gaa gct gaa gca gca gct aaa gaa gct ttg aaa aat gat	1788
Ser Tyr Glu Glu Ala Glu Ala Ala Lys Glu Ala Leu Lys Asn Asp	
485 490 495	
gat gta aat aaa tca tat act att aga caa ggt gca gat gga aga tat	1836
Asp Val Asn Lys Ser Tyr Thr Ile Arg Gln Gly Ala Asp Gly Arg Tyr	
500 505 510	
tac tat gta tta tca cca gta gaa gct gaa gaa gaa aaa cca gaa gca	1884
Tyr Tyr Val Leu Ser Pro Val Glu Ala Glu Glu Glu Lys Pro Glu Ala	
515 520 525	
caa aat gga tat gca aca tac gaa gaa gca gaa gca gca gct aag aaa	1932
Gln Asn Gly Tyr Ala Thr Tyr Glu Glu Ala Glu Ala Ala Lys Lys	
530 535 540	
gct ttg gaa aat gat cca atc aat aaa tct tac agc ata aga caa ggt	1980
Ala Leu Glu Asn Asp Pro Ile Asn Lys Ser Tyr Ser Ile Arg Gln Gly	
545 550 555 560	
gca gat gga aga tac tac tat gta tta tca cca gta gaa gct gaa act	2028
Ala Asp Gly Arg Tyr Tyr Tyr Val Leu Ser Pro Val Glu Ala Glu Thr	
565 570 575	
cct gaa aaa cca gta gaa cca tca gaa cca tca aca cca gat gta cca	2076
Pro Glu Lys Pro Val Glu Pro Ser Ser Glu Pro Ser Thr Pro Asp Val Pro	
580 585 590	
tca aac cca tca aat cca tca aca cca gat gtt cca tca act cct gat	2124
Ser Asn Pro Ser Asn Pro Ser Thr Pro Asp Val Pro Ser Thr Pro Asp	
595 600 605	
gta cca tca aat cct tca act cca gaa gtt cct tca aac cca tca act	2172
Val Pro Ser Asn Pro Ser Thr Pro Glu Val Pro Ser Asn Pro Ser Thr	
610 615 620	
cct ggt aat gaa gaa aaa cct ggt aac gaa caa aaa cct ggt aac gaa	2220
Pro Gly Asn Glu Glu Lys Pro Gly Asn Glu Gln Lys Pro Gly Asn Glu	
625 630 635 640	

caa aaa cct ggt aac gaa caa aaa cct ggt aat gaa caa aaa cct ggt Gln Lys Pro Gly Asn Glu Gln Lys Pro Gly Asn Glu Gln Lys Pro Gly 645 650 655	2268
aat gaa caa aaa cca gac caa cct tca aaa cca gaa aaa gaa gaa aat Asn Glu Gln Lys Pro Asp Gln Pro Ser Lys Pro Glu Lys Glu Glu Asn 660 665 670	2316
ggt aag ggt gga gta gat tct cca aag aaa aaa gaa aaa gct gca tta Gly Lys Gly Gly Val Asp Ser Pro Lys Lys Lys Glu Lys Ala Ala Leu 675 680 685	2364
cca aaa gct ggt agc gaa gct gaa atc tta aca tta gca gca gct tca Pro Lys Ala Gly Ser Glu Ala Glu Ile Leu Thr Leu Ala Ala Ala Ser 690 695 700	2412
tta tca agc gtt gca ggt gct ttc att tca ctt aaa aaa cgt aaa taa Leu Ser Ser Val Ala Gly Ala Phe Ile Ser Leu Lys Lys Arg Lys 705 710 715	2460
ttaatcttag ataaaaaata gattaatcta aaaaaatggg actttttataa gtcccatttt	2520
tgattgcaat gaaactgata caaaaaatgt atcagttttt tcattttacgc ttatttttctt	2580
gtgagtatgt ccaagtttgt cgtataatca tctatcaaaa gcctggcgag tttcgctttt	2640
tggattttgt cttcgaa	2657

<210> 2
 <211> 719
 <212> PRT
 <213> Peptostreptococcus magnus

<400> 2

Met Ala Ala Leu Ala Gly Ala Ile Val Val Thr Gly Gly Val Gly Ser 1 5 10 15
--

Tyr Ala Ala Asp Glu Pro Ile Asp Leu Glu Lys Leu Glu Glu Lys Arg 20 25 30

Asp Lys Glu Asn Val Gly Asn Leu Pro Lys Phe Asp Asn Glu Val Lys 35 40 45

Asp Gly Ser Glu Asn Pro Met Ala Lys Tyr Pro Asp Phe Asp Asp Glu 50 55 60

Ala Ser Thr Arg Phe Glu Thr Glu Asn Asn Glu Phe Glu Glu Lys Lys 65 70 75 80
--

Val Val Ser Asp Asn Phe Phe Asp Gln Ser Glu His Pro Phe Val Glu

85

90

95

Asn Lys Glu Glu Thr Pro Glu Thr Pro Glu Thr Asp Ser Glu Glu Glu
 100 105 110

Val Thr Ile Lys Ala Asn Leu Ile Phe Ala Asn Gly Ser Thr Gln Thr
 115 120 125

Ala Glu Phe Lys Gly Thr Phe Glu Lys Ala Thr Ser Glu Ala Tyr Ala
 130 135 140

Tyr Ala Asp Thr Leu Lys Lys Asp Asn Gly Glu Tyr Thr Val Asp Val
 145 150 155 160

Ala Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly Lys Glu Lys
 165 170 175

Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala Asn Leu Ile
 180 185 190

Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly Thr Phe Glu
 195 200 205

Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Ala Leu Lys Lys Asp
 210 215 220

Asn Gly Glu Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr Thr Leu Asn
 225 230 235 240

Ile Lys Phe Ala Gly Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu
 245 250 255

Val Thr Ile Lys Ala Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr
 260 265 270

Ala Glu Phe Lys Gly Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg
 275 280 285

Tyr Ala Asp Leu Leu Ala Lys Glu Asn Gly Lys Tyr Thr Val Asp Val
 290 295 300

Ala Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly Lys Glu Lys
 305 310 315 320

Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala Asn Leu Ile
 325 330 335

Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly Thr Phe Ala
 340 345 350

Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu Ala Lys Glu
 355 360 365

Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr Thr Ile Asn
 370 375 380

Ile Arg Phe Ala Gly Lys Lys Val Asp Glu Lys Pro Glu Glu Lys Glu
 385 390 395 400

Gln Val Thr Ile Lys Glu Asn Ile Tyr Phe Glu Asp Gly Thr Val Gln
 405 410 415

Thr Ala Thr Phe Lys Gly Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr
 420 425 430

Arg Tyr Ala Asp Leu Leu Ser Lys Glu His Gly Lys Tyr Thr Ala Asp
 435 440 445

Leu Glu Asp Gly Gly Tyr Thr Ile Asn Ile Arg Phe Ala Gly Lys Glu
 450 455 460

Glu Pro Glu Glu Thr Pro Glu Lys Pro Glu Val Gln Asp Gly Tyr Ala
 465 470 475 480

Ser Tyr Glu Glu Ala Glu Ala Ala Ala Lys Glu Ala Leu Lys Asn Asp
 485 490 495

Asp Val Asn Lys Ser Tyr Thr Ile Arg Gln Gly Ala Asp Gly Arg Tyr
 500 505 510

Tyr Tyr Val Leu Ser Pro Val Glu Ala Glu Glu Glu Lys Pro Glu Ala
 515 520 525

Gln Asn Gly Tyr Ala Thr Tyr Glu Glu Ala Glu Ala Ala Ala Lys Lys
 530 535 540

Ala Leu Glu Asn Asp Pro Ile Asn Lys Ser Tyr Ser Ile Arg Gln Gly
 545 550 555 560

Ala Asp Gly Arg Tyr Tyr Tyr Val Leu Ser Pro Val Glu Ala Glu Thr
 565 570 575

Pro Glu Lys Pro Val Glu Pro Ser Glu Pro Ser Thr Pro Asp Val Pro
 580 585 590

Ser Asn Pro Ser Asn Pro Ser Thr Pro Asp Val Pro Ser Thr Pro Asp
 595 600 605

Val Pro Ser Asn Pro Ser Thr Pro Glu Val Pro Ser Asn Pro Ser Thr
 610 615 620

Pro Gly Asn Glu Glu Lys Pro Gly Asn Glu Gln Lys Pro Gly Asn Glu
 625 630 635 640

Gln Lys Pro Gly Asn Glu Gln Lys Pro Gly Asn Glu Gln Lys Pro Gly
 645 650 655

Asn Glu Gln Lys Pro Asp Gln Pro Ser Lys Pro Glu Lys Glu Glu Asn
 660 665 670

Gly Lys Gly Gly Val Asp Ser Pro Lys Lys Lys Glu Lys Ala Ala Leu
 675 680 685

Pro Lys Ala Gly Ser Glu Ala Glu Ile Leu Thr Leu Ala Ala Ala Ser
 690 695 700

Leu Ser Ser Val Ala Gly Ala Phe Ile Ser Leu Lys Lys Arg Lys
 705 710 715

<210> 3
 <211> 228
 <212> DNA
 <213> Peptostreptococcus magnus

<220>
 <221> CDS
 <222> (1)..(228)
 <223> PpL1

<400> 3

<223> PpL2

<400> 5

aaa	gaa	aaa	aca	cca	gaa	gaa	cca	aaa	gaa	gaa	gtt	act	att	aaa	gca	48
Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala	
1				5				10						15		

aac	tta	atc	tat	gca	gat	gga	aaa	aca	caa	aca	gca	gaa	ttc	aaa	gga	96
Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	
		20					25						30			

aca	ttt	gaa	gaa	gca	aca	gca	gaa	gca	tac	aga	tat	gca	gat	gca	tta	144
Thr	Phe	Glu	Glu	Ala	Thr	Ala	Glu	Ala	Tyr	Arg	Tyr	Ala	Asp	Ala	Leu	
		35					40					45				

aag	aag	gac	aat	gga	gaa	tat	aca	gta	gac	gtt	gca	gat	aaa	ggt	tat	192
Lys	Lys	Asp	Asn	Gly	Glu	Tyr	Thr	Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr	
	50					55					60					

act	tta	aat	att	aaa	ttt	gct	gga									216
Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly									
65						70										

<210> 6

<211> 72

<212> PRT

<213> Peptostreptococcus magnus

<400> 6

Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala
1				5				10						15	

Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly
		20					25						30		

Thr	Phe	Glu	Glu	Ala	Thr	Ala	Glu	Ala	Tyr	Arg	Tyr	Ala	Asp	Ala	Leu
		35					40					45			

Lys	Lys	Asp	Asn	Gly	Glu	Tyr	Thr	Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr
	50					55					60				

Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly								
65						70									

<210> 7

<211> 216

<212> DNA

<213> Peptostreptococcus magnus

<220>
 <221> CDS
 <222> (1)..(216)
 <223> PpL3

 <400> 7
 aaa gaa aaa aca cca gaa gaa cca aaa gaa gaa gtt act att aaa gca 48
 Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala
 1 5 10 15

 aac tta atc tat gca gat gga aaa aca caa aca gca gaa ttc aaa gga 96
 Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly
 20 25 30

 aca ttt gaa gaa gca aca gca gaa gca tac aga tat gct gac tta tta 144
 Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu
 35 40 45

 gca aaa gaa aat ggt aaa tat aca gta gac gtt gca gat aaa ggt tat 192
 Ala Lys Glu Asn Gly Lys Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr
 50 55 60

 act tta aat att aaa ttt gct gga 216
 Thr Leu Asn Ile Lys Phe Ala Gly
 65 70

 <210> 8
 <211> 72
 <212> PRT
 <213> Peptostreptococcus magnus

 <400> 8

 Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala
 1 5 10 15

 Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly
 20 25 30

 Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu
 35 40 45

 Ala Lys Glu Asn Gly Lys Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr
 50 55 60

 Thr Leu Asn Ile Lys Phe Ala Gly
 65 70

 <210> 9
 <211> 216
 <212> DNA

<213> Peptostreptococcus magnus

<220>

<221> CDS

<222> (1)..(216)

<223> PpL4

<400> 9

aaa gaa aaa aca cca gaa gaa cca aaa gaa gaa gtt act att aaa gca	48
Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala	
1 5 10 15	

aac tta atc tat gca gat gga aaa act caa aca gca gag ttc aaa gga	96
Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly	
20 25 30	

aca ttt gca gaa gca aca gca gaa gca tac aga tac gct gac tta tta	144
Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu	
35 40 45	

gca aaa gaa aat ggt aaa tat aca gca gac tta gaa gat ggt gga tac	192
Ala Lys Glu Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr	
50 55 60	

act att aat att aga ttt gca ggt	216
Thr Ile Asn Ile Arg Phe Ala Gly	
65 70	

<210> 10

<211> 72

<212> PRT

<213> Peptostreptococcus magnus

<400> 10

Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala	
1 5 10 15	

Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly	
20 25 30	

Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu	
35 40 45	

Ala Lys Glu Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr	
50 55 60	

Thr Ile Asn Ile Arg Phe Ala Gly	
65 70	

<210> 11
 <211> 219
 <212> DNA
 <213> Peptostreptococcus magnus

<220>
 <221> CDS
 <222> (1)..(219)
 <223> PpL5

<400> 11
 aag aaa gtt gac gaa aaa cca gaa gaa aaa gaa caa gta aca att aaa 48
 Lys Lys Val Asp Glu Lys Pro Glu Glu Lys Glu Gln Val Thr Ile Lys
 1 5 10 15
 gaa aat ata tat ttt gaa gat gga aca gta caa act gca aca ttt aaa 96
 Glu Asn Ile Tyr Phe Glu Asp Gly Thr Val Gln Thr Ala Thr Phe Lys
 20 25 30
 gga aca ttt gca gaa gcg aca gca gaa gca tac aga tat gca gat ttg 144
 Gly Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu
 35 40 45
 tta tca aaa gaa cat ggt aaa tac aca gca gac ttg gaa gat ggt gga 192
 Leu Ser Lys Glu His Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly
 50 55 60
 tac act atc aac att aga ttt gct gga 219
 Tyr Thr Ile Asn Ile Arg Phe Ala Gly
 65 70

<210> 12
 <211> 73
 <212> PRT
 <213> Peptostreptococcus magnus

<400> 12
 Lys Lys Val Asp Glu Lys Pro Glu Glu Lys Glu Gln Val Thr Ile Lys
 1 5 10 15
 Glu Asn Ile Tyr Phe Glu Asp Gly Thr Val Gln Thr Ala Thr Phe Lys
 20 25 30
 Gly Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu
 35 40 45
 Leu Ser Lys Glu His Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly
 50 55 60
 Tyr Thr Ile Asn Ile Arg Phe Ala Gly

<210> 13
 <211> 1269
 <212> DNA
 <213> Sindbis virus

<220>
 <221> misc_feature
 <222> (1)..(1269)
 <223> E2 potein gene from strain AR339

<220>
 <221> CDS
 <222> (1)..(1269)
 <223> E2 glycoprotein

<220>
 <221> variation
 <222> (3)..(3)
 <223> a in AR339; c in HRsp

<220>
 <221> variation
 <222> (8)..(8)
 <223> c in AR339; t in HRsp

<220>
 <221> variation
 <222> (68)..(68)
 <223> a in AR339; t in HRsp

<220>
 <221> variation
 <222> (208)..(208)
 <223> g in AR339; a in HRsp

<220>
 <221> variation
 <222> (514)..(514)
 <223> g in AR339; a in HRsp

<400> 13
 aga gtc act gac gac ttt acc ctg acc agc ccc tac ttg ggc aca tgc 48
 Arg Val Thr Asp Asp Phe Thr Leu Thr Ser Pro Tyr Leu Gly Thr Cys
 1 5 10 15
 tcg tac tgc cac cat act gaa ccg tgc ttc agc cct gtt aag atc gag 96
 Ser Tyr Cys His His Thr Glu Pro Cys Phe Ser Pro Val Lys Ile Glu
 20 25 30
 cag gtc tgg gac gaa gcg gac gat aac acc ata cgc ata cag act tcc 144
 Gln Val Trp Asp Glu Ala Asp Asp Asn Thr Ile Arg Ile Gln Thr Ser
 35 40 45

gcc cag ttt gga tac gac caa agc gga gca gca agc gca aac aag tac Ala Gln Phe Gly Tyr Asp Gln Ser Gly Ala Ala Ser Ala Asn Lys Tyr 50 55 60	192
cgc tac atg tcg ctt gag cag gat cac acc gtt aaa gaa ggc acc atg Arg Tyr Met Ser Leu Glu Gln Asp His Thr Val Lys Glu Gly Thr Met 65 70 75 80	240
gat gac atc aag att agc acc tca gga ccg tgt aga agg ctt agc tac Asp Asp Ile Lys Ile Ser Thr Ser Gly Pro Cys Arg Arg Leu Ser Tyr 85 90 95	288
aaa gga tac ttt ctc ctc gca aaa tgc cct cca ggg gac agc gta acg Lys Gly Tyr Phe Leu Leu Ala Lys Cys Pro Pro Gly Asp Ser Val Thr 100 105 110	336
gtt agc ata gtg agt agc aac tca gca acg tca tgt aca ctg gcc cgc Val Ser Ile Val Ser Ser Asn Ser Ala Thr Ser Cys Thr Leu Ala Arg 115 120 125	384
aag ata aaa cca aaa ttc gtg gga cgg gaa aaa tat gat cta cct ccc Lys Ile Lys Pro Lys Phe Val Gly Arg Glu Lys Tyr Asp Leu Pro Pro 130 135 140	432
gtt cac ggt aaa aaa att cct tgc aca gtg tac gac cgt ctg aaa gaa Val His Gly Lys Lys Ile Pro Cys Thr Val Tyr Asp Arg Leu Lys Glu 145 150 155 160	480
aca act gca ggc tac atc act atg cac agg ccg gga ccg cac gct tat Thr Thr Ala Gly Tyr Ile Thr Met His Arg Pro Gly Pro His Ala Tyr 165 170 175	528
aca tcc tac ctg gaa gaa tca tca ggg aaa gtt tac gca aag ccg cca Thr Ser Tyr Leu Glu Glu Ser Ser Gly Lys Val Tyr Ala Lys Pro Pro 180 185 190	576
tct ggg aag aac att acg tat gag tgc aag tgc ggc gac tac aag acc Ser Gly Lys Asn Ile Thr Tyr Glu Cys Lys Cys Gly Asp Tyr Lys Thr 195 200 205	624
gga acc gtt tcg acc cgc acc gaa atc act ggt tgc acc gcc atc aag Gly Thr Val Ser Thr Arg Thr Glu Ile Thr Gly Cys Thr Ala Ile Lys 210 215 220	672
cag tgc gtc gcc tat aag agc gac caa acg aag tgg gtc ttc aac tca Gln Cys Val Ala Tyr Lys Ser Asp Gln Thr Lys Trp Val Phe Asn Ser 225 230 235 240	720
ccg gac ttg atc aga cat gac gac cac acg gcc caa ggg aaa ttg cat Pro Asp Leu Ile Arg His Asp Asp His Thr Ala Gln Gly Lys Leu His 245 250 255	768
ttg cct ttc aag ttg atc ccg agt acc tgc atg gtc cct gtt gcc cac Leu Pro Phe Lys Leu Ile Pro Ser Thr Cys Met Val Pro Val Ala His 260 265 270	816
gcg ccg aat gta ata cat ggc ttt aaa cac atc agc ctc caa tta gat	864

Ala	Pro	Asn	Val	Ile	His	Gly	Phe	Lys	His	Ile	Ser	Leu	Gln	Leu	Asp		
		275					280					285					
aca	gac	cac	ttg	aca	ttg	ctc	acc	acc	agg	aga	cta	ggg	gca	aac	ccg		912
Thr	Asp	His	Leu	Thr	Leu	Leu	Thr	Thr	Arg	Arg	Leu	Gly	Ala	Asn	Pro		
	290					295					300						
gaa	cca	acc	act	gaa	tgg	atc	gtc	gga	aag	acg	gtc	aga	aac	ttc	acc		960
Glu	Pro	Thr	Thr	Glu	Trp	Ile	Val	Gly	Lys	Thr	Val	Arg	Asn	Phe	Thr		
305					310					315					320		
gtc	gac	cga	gat	ggc	ctg	gaa	tac	ata	tgg	gga	aat	cat	gag	cca	gtg		1008
Val	Asp	Arg	Asp	Gly	Leu	Glu	Tyr	Ile	Trp	Gly	Asn	His	Glu	Pro	Val		
				325					330					335			
agg	gtc	tat	gcc	caa	gag	tca	gca	cca	gga	gac	cct	cac	gga	tgg	cca		1056
Arg	Val	Tyr	Ala	Gln	Glu	Ser	Ala	Pro	Gly	Asp	Pro	His	Gly	Trp	Pro		
			340					345					350				
cac	gaa	ata	gta	cag	cat	tac	tac	cat	cgc	cat	cct	gtg	tac	acc	atc		1104
His	Glu	Ile	Val	Gln	His	Tyr	Tyr	His	Arg	His	Pro	Val	Tyr	Thr	Ile		
		355					360					365					
tta	gcc	gtc	gca	tca	gct	acc	gtg	gcg	atg	atg	att	ggc	gta	acc	gtt		1152
Leu	Ala	Val	Ala	Ser	Ala	Thr	Val	Ala	Met	Met	Ile	Gly	Val	Thr	Val		
	370					375					380						
gca	gtg	tta	tgt	gcc	tgt	aaa	gcg	cgc	cgt	gag	tgc	ctg	acg	cca	tac		1200
Ala	Val	Leu	Cys	Ala	Cys	Lys	Ala	Arg	Arg	Glu	Cys	Leu	Thr	Pro	Tyr		
385					390					395					400		
gcc	ctg	gcc	cca	aac	gcc	gta	atc	cca	act	tcg	ctg	gca	ctc	ttg	tgc		1248
Ala	Leu	Ala	Pro	Asn	Ala	Val	Ile	Pro	Thr	Ser	Leu	Ala	Leu	Leu	Cys		
				405					410					415			
tgc	gtt	agg	tcg	gcc	aat	gct											1269
Cys	Val	Arg	Ser	Ala	Asn	Ala											
			420														

<210> 14
 <211> 423
 <212> PRT
 <213> Sindbis virus

<400> 14

Arg	Val	Thr	Asp	Asp	Phe	Thr	Leu	Thr	Ser	Pro	Tyr	Leu	Gly	Thr	Cys		
1				5					10					15			

Ser	Tyr	Cys	His	His	Thr	Glu	Pro	Cys	Phe	Ser	Pro	Val	Lys	Ile	Glu		
			20					25					30				

Gln	Val	Trp	Asp	Glu	Ala	Asp	Asp	Asn	Thr	Ile	Arg	Ile	Gln	Thr	Ser		
		35					40					45					

Ala	Gln	Phe	Gly	Tyr	Asp	Gln	Ser	Gly	Ala	Ala	Ser	Ala	Asn	Lys	Tyr	50	55	60	
Arg	Tyr	Met	Ser	Leu	Glu	Gln	Asp	His	Thr	Val	Lys	Glu	Gly	Thr	Met	65	70	75	80
Asp	Asp	Ile	Lys	Ile	Ser	Thr	Ser	Gly	Pro	Cys	Arg	Arg	Leu	Ser	Tyr	85	90	95	
Lys	Gly	Tyr	Phe	Leu	Leu	Ala	Lys	Cys	Pro	Pro	Gly	Asp	Ser	Val	Thr	100	105	110	
Val	Ser	Ile	Val	Ser	Ser	Asn	Ser	Ala	Thr	Ser	Cys	Thr	Leu	Ala	Arg	115	120	125	
Lys	Ile	Lys	Pro	Lys	Phe	Val	Gly	Arg	Glu	Lys	Tyr	Asp	Leu	Pro	Pro	130	135	140	
Val	His	Gly	Lys	Lys	Ile	Pro	Cys	Thr	Val	Tyr	Asp	Arg	Leu	Lys	Glu	145	150	155	160
Thr	Thr	Ala	Gly	Tyr	Ile	Thr	Met	His	Arg	Pro	Gly	Pro	His	Ala	Tyr	165	170	175	
Thr	Ser	Tyr	Leu	Glu	Glu	Ser	Ser	Gly	Lys	Val	Tyr	Ala	Lys	Pro	Pro	180	185	190	
Ser	Gly	Lys	Asn	Ile	Thr	Tyr	Glu	Cys	Lys	Cys	Gly	Asp	Tyr	Lys	Thr	195	200	205	
Gly	Thr	Val	Ser	Thr	Arg	Thr	Glu	Ile	Thr	Gly	Cys	Thr	Ala	Ile	Lys	210	215	220	
Gln	Cys	Val	Ala	Tyr	Lys	Ser	Asp	Gln	Thr	Lys	Trp	Val	Phe	Asn	Ser	225	230	235	240
Pro	Asp	Leu	Ile	Arg	His	Asp	Asp	His	Thr	Ala	Gln	Gly	Lys	Leu	His	245	250	255	
Leu	Pro	Phe	Lys	Leu	Ile	Pro	Ser	Thr	Cys	Met	Val	Pro	Val	Ala	His	260	265	270	

Ala Pro Asn Val Ile His Gly Phe Lys His Ile Ser Leu Gln Leu Asp
 275 280 285

Thr Asp His Leu Thr Leu Leu Thr Thr Arg Arg Leu Gly Ala Asn Pro
 290 295 300

Glu Pro Thr Thr Glu Trp Ile Val Gly Lys Thr Val Arg Asn Phe Thr
 305 310 315 320

Val Asp Arg Asp Gly Leu Glu Tyr Ile Trp Gly Asn His Glu Pro Val
 325 330 335

Arg Val Tyr Ala Gln Glu Ser Ala Pro Gly Asp Pro His Gly Trp Pro
 340 345 350

His Glu Ile Val Gln His Tyr Tyr His Arg His Pro Val Tyr Thr Ile
 355 360 365

Leu Ala Val Ala Ser Ala Thr Val Ala Met Met Ile Gly Val Thr Val
 370 375 380

Ala Val Leu Cys Ala Cys Lys Ala Arg Arg Glu Cys Leu Thr Pro Tyr
 385 390 395 400

Ala Leu Ala Pro Asn Ala Val Ile Pro Thr Ser Leu Ala Leu Leu Cys
 405 410 415

Cys Val Arg Ser Ala Asn Ala
 420

<210> 15
 <211> 11703
 <212> DNA
 <213> Sindbis virus

<220>
 <221> misc_feature
 <222> (1)..(11703)
 <223> Sindbis virus HRsp and wild-type strains complete genome

<220>
 <221> CDS
 <222> (8439)..(8630)
 <223> E3 protein

<220>
 <221> CDS
 <222> (8631)..(9899)
 <223> E2 protein

<400> 15
 attgacggcg tagtacacac tattgaatca aacagccgac caattgcact accatcacia 60
 tggagaagcc agtagtaaac gtagacgtag acccccagag tccgtttgtc gtgcaactgc 120
 aaaaaagctt cccgcaattt gaggtagtag cacagcaggt cactccaaat gaccatgcta 180
 atgccagagc attttcgcac ctggccagta aactaatcga gctggagggt cctaccacag 240
 cgacgatctt ggacataggc agcgcaccgg ctcgtagaat gttttccgag caccagtatc 300
 attgtgtctg ccccatgcgt agtccagaag acccggaccg catgatgaaa tacgccagta 360
 aactggcggg aaaagcgtgc aagattacaa acaagaactt gcatgagaag attaaggatc 420
 tccggaccgt acttgatacg ccggatgctg aaacaccatc gctctgcttt cacaacgatg 480
 ttacctgcaa catgcgtgcc gaatatccg tcatgcagga cgtgtatatc aacgctcccg 540
 gaactatcta tcatcaggct atgaaaggcg tgcggaccct gtactggatt ggcttcgaca 600
 ccacccagtt catgttctcg gctatggcag gttcgtaccc tgcgtacaac accaactggg 660
 ccgacgagaa agtccttgaa gcgcgtaaca tcggactttg cagcaciaag ctgagtgaag 720
 gtaggacagg aaaattgtcg ataattgagga agaaggaggt gaagcccggg tcgcggggtt 780
 atttctccgt aggatcgaca ctttatccag aacacagagc cagcttgacg agctggcatc 840
 ttccatcggg gttccacttg aatggaaagc agtcgtacac ttgccgctgt gatacagtgg 900
 tgagttgcga aggctacgta gtgaagaaaa tcaccatcag tcccgggatc acgggagaaa 960
 ccgtgggata cgcggttaca cacaatagcg agggcttctt gctatgcaaa gttactgaca 1020
 cagtaaaagg agaacgggta tcgttccctg tgtgcacgta catcccggcc accatatgcg 1080
 atcagatgac tgggtataatg gccacggata tatcacctga cgatgcacia aaacttctgg 1140
 ttgggctcaa ccagcgaatt gtcattaacg gtaggactaa caggaacacc aacaccatgc 1200
 aaaattacct tctgccgatc atagcacaag ggttcagcaa atgggctaag gagcgcaagg 1260
 atgatcttga taacgagaaa atgctgggta ctagagaacg caagcttacg tatggctgct 1320
 tgtgggctgt tcgcactaag aaagtacatt cgttttatcg cccacctgga acgcagacct 1380
 gcgtaaaagt cccagcctct tttagcgctt ttcccatgtc gtccgtatgg acgacctctt 1440
 tgcccatgtc gctgaggcag aaattgaaac tggcattgca accaaagaag gaggaaaaac 1500
 tgctgcaggt ctccggaggaa ttagtcatgg agccaaggc tgcttttgag gatgctcagg 1560

aggaagccag agcggagaag ctccgagaag cacttccacc attagtggca gacaaaggca	1620
tcgaggcagc cgcagaagtt gtctgcgaag tggaggggct ccaggcggac atcggagcag	1680
cattagttga aaccccgcg cgtcacgtaa ggataatacc tcaagcaaat gaccgtatga	1740
tcggacagta tatcgttgct tcgccaaact ctgtgctgaa gaatgccaaa ctgcgaccag	1800
cgcacccgct agcagatcag gttaagatca taacacactc cggaagatca ggaaggtagc	1860
cggtcgaacc atacgacgct aaagtactga tgccagcagg aggtgccgta ccatggccag	1920
aattcctagc actgagttag agcgccacgt tagtgtacaa cgaaagagag tttgtgaacc	1980
gcaaactata ccacattgcc atgcatggcc ccgccaagaa tacagaagag gagcagtaca	2040
aggttacaaa ggcagagctt gcagaaacag agtacgtgtt tgacgtggac aagaagcgtt	2100
gcgttaagaa ggaagaagcc tcaggtctgg tcctctcggg agaactgacc aaccctccct	2160
atcatgagct agctctggag ggactgaaga cccgacctgc ggtcccgtac aaggtcgaaa	2220
caataggagt gataggcaca ccgggggtcgg gcaagtcagc tattatcaag tcaactgtca	2280
cggcacgaga tcttgttacc agcggaaaga aagaaaattg tcgcgaaatt gaggccgacg	2340
tgctaagact gaggggtatg cagattacgt cgaagacagt agattcgggt atgctcaacg	2400
gatgccacaa agccgtagaa gtgctgtacg ttgacgaagc gttcgcgtgc cacgcaggag	2460
cactacttgc cttgattgct atcgtcaggc cccgcaagaa ggtagtacta tgcggagacc	2520
ccatgcaatg cggattcttc aacatgatgc aactaaaggt acatttcaat caccctgaaa	2580
aagacatatg caccaagaca ttctacaagt atatctcccg gcgttgcaca cagccagtta	2640
cagctattgt atcgacactg cattacgatg gaaagatgaa aaccacgaac ccgtgcaaga	2700
agaacattga aatcgatatt acagggggcca caaagccgaa gccaggggat atcatcctga	2760
catgtttccg cgggtggggt aagcaattgc aaatcgacta tcccggacat gaagtaatga	2820
cagccgcggc ctcaaaaggg ctaaccagaa aaggagtgtg tgccgtccgg caaaaagtca	2880
atgaaaaccc actgtacgcg atcacatcag agcatgtgaa cgtgttgctc acccgactg	2940
aggacaggct agtgtggaaa accttgcagg gcgacctatg gattaagcag cccactaaca	3000
tacctaaagg aaactttcag gctactatag aggactggga agctgaacac aagggaataa	3060
ttgctgcaat aaacagcccc actccccgtg ccaatccgtt cagctgcaag accaacgttt	3120
gctgggcgaa agcattggaa ccgatactag ccacggccgg tatcgtactt accggttgcc	3180
agtggagcga actgttccca cagtttgcg atgacaaacc acattcggcc atttacgcct	3240

tagacgtaat	ttgcattaag	tttttcggca	tggacttgac	aagcggactg	ttttctaaac	3300
agagcatccc	actaacgtac	catccccgcc	attcagcgag	gccggtagct	cattgggaca	3360
acagcccagg	aacccgcaag	tatgggtacg	atcacgccat	tgccgccgaa	ctctcccgtc	3420
gatttccggt	gttccagcta	gctgggaagg	gcacacaact	tgatttgcag	acggggagaa	3480
ccagagttat	ctctgcacag	cataacctgg	tcccggtgaa	ccgcaatctt	cctcacgcct	3540
tagtccccga	gtacaaggag	aagcaacccg	gcccgggtcaa	aaaattcttg	aaccagttca	3600
aacaccactc	agtacttggt	gtatcagagg	aaaaaattga	agctccccgt	aagagaatcg	3660
aatggatcgc	cccgattggc	atagccgggt	cagataagaa	ctacaacctg	gctttcgggt	3720
ttccgccgca	ggcacgggtac	gacctgggtg	tcataacat	tggaactaaa	tacagaaacc	3780
accactttca	gcagtgcgaa	gacatgcgg	cgaccttaaa	aaccctttcg	cgttcggccc	3840
tgaattgcct	taaccagga	ggcaccctcg	tgggtgaagtc	ctatgggtac	gccgaccgca	3900
acagtgagga	cgtagtcacc	gctcttgcca	gaaagtttgt	cagggtgtct	gcagcgagac	3960
cagattgtgt	ctcaagcaat	acagaaatgt	acctgatttt	ccgacaacta	gacaacagcc	4020
gtacacggca	attcaccccc	caccatctga	attgcgtgat	ttcgtcctgt	tatgagggta	4080
caagagatgg	agttggagcc	gcgccgtcat	accgcaccaa	aaggagagaat	attgctgact	4140
gtcaagagga	agcagttgtc	aacgcagcca	atccgctggg	tagaccaggc	gaaggagtct	4200
gccgtgccat	ctataaacgt	tggccgacca	gtttttaccga	ttcagccacg	gagacaggca	4260
ccgcaagaat	gactgtgtgc	ctaggaaaga	aagtgatcca	cgcggtcggc	cctgatttcc	4320
ggaagcacc	agaagcagaa	gccttgaaat	tgctacaaaa	cgcctaccat	gcagtggcag	4380
acttagtaaa	tgaacataac	atcaagtctg	tcgccatttc	actgctatct	acaggcattt	4440
acgcagccgg	aaaagaccgc	cttgaagtat	cacttaactg	cttgacaacc	gcgctagaca	4500
gaactgacgc	ggacgtaacc	atctattgcc	tggataagaa	gtggaaggaa	agaatcgacg	4560
cggcactcca	acttaaggag	tctgtaacag	agctgaagga	tgaagatatg	gagatcgacg	4620
atgagttagt	atggatccat	ccagacagtt	gcttgaaggg	aagaaaggga	ttcagtacta	4680
caaaaggaaa	attgtattcg	tacttcgaag	gcaccaaatt	ccatcaagca	gcaaaagaca	4740
tggcggagat	aaaggtcctg	ttccctaatt	accaggaaag	taatgaacaa	ctgtgtgcct	4800
acatatggg	tgagaccatg	gaagcaatcc	gcgaaaagtg	cccggtcgac	cataaccctg	4860
cgtctagccc	gccccaaacg	ttgccgtgcc	tttgcattga	tgccatgacg	ccagaaaggg	4920
tccacagact	tagaagcaat	aacgtcaaag	aagttacagt	atgctcctcc	accccccttc	4980

ctāagcacia	aattaagaat	gttcagaagg	ttcagtgcac	gaaagtagtc	ctgtttaatc	5040
cgcacactcc	cgcattcggt	cccgcccgtg	agtacataga	agtgccagaa	cagcctaccg	5100
ctcctcctgc	acaggccgag	gaggcccccg	aagttgtagc	gacaccgtca	ccatctacag	5160
ctgataacac	ctcgcttgat	gtcacagaca	tctcactgga	tatggatgac	agtagcgaag	5220
gctcactttt	ttcgagcttt	agcggatcgg	acaactctat	tactagtatg	gacagttggt	5280
cgtcaggacc	tagttcacta	gagatagtag	accgaaggca	ggtggtggtg	gctgacgttc	5340
atgccgtcca	agagcctgcc	cctattccac	cgccaaggct	aaagaagatg	gcccgcctgg	5400
cagcggcaag	aaaagagccc	actccaccgg	caagcaatag	ctctgagtc	ctccacctct	5460
cttttggtgg	ggtatccatg	tccctcggat	caattttcga	cggagagacg	gcccgccagg	5520
cagcgggtaca	acccttgcca	acaggcccca	cggatgtgcc	tatgtctttc	ggatcgtttt	5580
ccgacggaga	gattgatgag	ctgagccgca	gagtaactga	gtccgaaccc	gtcctgtttg	5640
gatcatttga	accgggcgaa	gtgaactcaa	ttatatcgtc	ccgatcagcc	gtatcttttc	5700
cactacgcaa	gcagagacgt	agacgcagga	gcaggaggac	tgaatactga	ctaaccgggg	5760
taggtgggta	catattttcg	acggacacag	gccctgggca	cttgcaaaag	aagtccgttc	5820
tgcagaacca	gcttacagaa	ccgaccttgg	agcgcaatgt	cctggaaaga	attcatgccc	5880
cgggtgctcga	cacgtcgaaa	gaggaacaac	tcaaactcag	gtaccagatg	atgcccaccg	5940
aagccaacaa	aagtaggtac	cagtctcgta	aagtagaaaa	tcagaaagcc	ataaccactg	6000
agcgactact	gtcaggacta	cgactgtata	actctgccac	agatcagcca	gaatgctata	6060
agatcaccta	tccgaaacca	ttgtactcca	gtagcgtacc	ggcgaactac	tccgatccac	6120
agttcgctgt	agctgtctgt	aacaactatc	tgcattgagaa	ctatccgaca	gtagcatctt	6180
atcagattac	tgacgagtac	gatgcttact	tggatatggt	agacggggaca	gtcgccctgcc	6240
tggatactgc	aaccttctgc	cccgctaagc	ttagaagtta	cccgaaaaaa	catgagtata	6300
gagccccgaa	tatccgcagt	gcggttccat	cagcgatgca	gaacacgcta	caaatgtgc	6360
tcattgccgc	aactaaaaga	aattgcaacg	tcacgcagat	gcgtgaactg	ccaacactgg	6420
actcagcgac	attcaatgtc	gaatgctttc	gaaaatatgc	atgtaatgac	gagtattggg	6480
aggagttcgc	tcggaagcca	attaggatta	ccactgagtt	tgtcaccgca	tatgtagcta	6540
gactgaaagg	ccctaaggcc	gccgcactat	ttgcaaagac	gtataatttg	gtcccattgc	6600
aagaagtgcc	tatggataga	ttcgtcatgg	acatgaaaag	agacgtgaaa	gttacaccag	6660

gcacgaaaca	cacagaagaa	agaccgaaag	tacaagtgat	acaagccgca	gaacccctgg	6720
cgactgctta	cttatgcggg	attcaccggg	aattagtgcg	taggcttacg	gccgtcttgc	6780
ttccaaacat	tcacacgctt	tttgacatgt	cggcggagga	ttttgatgca	atcatagcag	6840
aacacttcaa	gcaaggcgac	ccggtactgg	agacggatat	cgcattcatc	gacaaaagcc	6900
aagacgacgc	tatggcgctta	accggtctga	tgatcttgga	ggacctgggt	gtggatcaac	6960
cactactcga	cttgatcgag	tgcgcccttg	gagaaatata	atccacccat	ctacctacgg	7020
gtactcgttt	taaattcggg	gcgatgatga	aatccggaat	gttcctcaca	ctttttgtca	7080
acacagtttt	gaatgtcgtt	atcgccagca	gagtactaga	agagcggctt	aaaacgtcca	7140
gatgtgcagc	gttcattggc	gacgacaaca	tcatacatgg	agtagtatct	gacaaagaaa	7200
tggctgagag	gtgcgccacc	tggctcaaca	tggagggttaa	gatcatcgac	gcagtcacgc	7260
gtgagagacc	accttacttc	tgcggcggat	ttatcttgca	agattcgggt	acttccacag	7320
cgtgccgcgt	ggcggatccc	ctgaaaaggc	tgtttaagtt	gggtaaaccg	ctcccagccg	7380
acgacgagca	agacgaagac	agaagacgcg	ctctgctaga	tgaaacaaag	gcgtgggttta	7440
gagtaggtat	aacaggcact	ttagcagtgg	ccgtgacgac	ccggtatgag	gtagacaata	7500
ttacacctgt	cctactggca	ttgagaactt	ttgcccagag	caaaagagca	ttccaagcca	7560
tcagagggga	aataaagcat	ctctacgggtg	gtcctaaata	gtcagcatag	tacatttcat	7620
ctgactaata	ctacaacacc	accaccatga	atagaggatt	ctttaacatg	ctcggccgcc	7680
gccccttccc	ggccccact	gccatgtgga	ggccgcggag	aaggaggcag	gcggccccga	7740
tgctgccccg	caacgggctg	gcttctcaaa	tccagcaact	gaccacagcc	gtcagtgccc	7800
tagtcattgg	acaggcaact	agacctcaac	ccccacgtcc	acgcccgcga	ccgcgccaga	7860
agaagcaggc	gccaagcaa	ccaccgaagc	cgaagaaacc	aaaaacgcag	gagaagaaga	7920
agaagcaacc	tgcaaaaccc	aaaccgcgaa	agagacagcg	catggcactt	aagttggagg	7980
ccgacagatt	gttcgacgtc	aagaacgagg	acggagatgt	catcgggcac	gcactggcca	8040
tggaaggaaa	ggtaatgaaa	cctctgcacg	tgaaaggaa	catcgaccac	cctgtgctat	8100
caaagctcaa	atttaccaag	tcgtcagcat	acgacatgga	gttcgcacag	ttgccagtca	8160
acatgagaag	tgaggcattc	acctacacca	gtgaacaccc	cgaaggattc	tataactggc	8220
accacggagc	ggtgcagtat	agtggaggta	gatttaccat	ccctcgcgga	gtaggaggca	8280
gaggagacag	cggtcgtccg	atcatggata	actccggtcg	ggttgtcgcg	atagtcctcg	8340
gtggcgctga	tgaaggaa	cgaactgccc	tttcggtcgt	cacctggaat	agtaaaggga	8400

agacaattaa gacgacccccg gaagggacag aagagtgg tcc gca gca cca ctg gtc	8456
Ser Ala Ala Pro Leu Val	
1 5	
acg gca atg tgt ttg ctc gga aat gtg agc ttc cca tgc gac cgc ccg	8504
Thr Ala Met Cys Leu Leu Gly Asn Val Ser Phe Pro Cys Asp Arg Pro	
10 15 20	
ccc aca tgc tat acc cgc gaa cct tcc aga gcc ctc gac atc ctt gaa	8552
Pro Thr Cys Tyr Thr Arg Glu Pro Ser Arg Ala Leu Asp Ile Leu Glu	
25 30 35	
gag aac gtg aac cat gag gcc tac gat acc ctg ctc aat gcc ata ttg	8600
Glu Asn Val Asn His Glu Ala Tyr Asp Thr Leu Leu Asn Ala Ile Leu	
40 45 50	
cgg tgc gga tcg tct ggc aga agc aaa aga agc gtc att gac gac ttt	8648
Arg Cys Gly Ser Ser Gly Arg Ser Lys Arg Ser Val Ile Asp Asp Phe	
55 60 65 70	
acc ctg acc agc ccc tac ttg ggc aca tgc tcg tac tgc cac cat act	8696
Thr Leu Thr Ser Pro Tyr Leu Gly Thr Cys Ser Tyr Cys His His Thr	
75 80 85	
gta ccg tgc ttc agc cct gtt aag atc gag cag gtc tgg gac gaa gcg	8744
Val Pro Cys Phe Ser Pro Val Lys Ile Glu Gln Val Trp Asp Glu Ala	
90 95 100	
gac gat aac acc ata cgc ata cag act tcc gcc cag ttt gga tac gac	8792
Asp Asp Asn Thr Ile Arg Ile Gln Thr Ser Ala Gln Phe Gly Tyr Asp	
105 110 115	
caa agc gga gca gca agc gca aac aag tac cgc tac atg tcg ctt aag	8840
Gln Ser Gly Ala Ala Ser Ala Asn Lys Tyr Arg Tyr Met Ser Leu Lys	
120 125 130	
cag gat cac acc gtt aaa gaa ggc acc atg gat gac atc aag att agc	8888
Gln Asp His Thr Val Lys Glu Gly Thr Met Asp Asp Ile Lys Ile Ser	
135 140 145 150	
acc tca gga ccg tgt aga agg ctt agc tac aaa gga tac ttt ctc ctc	8936
Thr Ser Gly Pro Cys Arg Arg Leu Ser Tyr Lys Gly Tyr Phe Leu Leu	
155 160 165	
gca aaa tgc cct cca ggg gac agc gta acg gtt agc ata gtg agt agc	8984
Ala Lys Cys Pro Pro Gly Asp Ser Val Thr Val Ser Ile Val Ser Ser	
170 175 180	
aac tca gca acg tca tgt aca ctg gcc cgc aag ata aaa cca aaa ttc	9032
Asn Ser Ala Thr Ser Cys Thr Leu Ala Arg Lys Ile Lys Pro Lys Phe	
185 190 195	
gtg gga cgg gaa aaa tat gat cta cct ccc gtt cac ggt aaa aaa att	9080
Val Gly Arg Glu Lys Tyr Asp Leu Pro Pro Val His Gly Lys Lys Ile	
200 205 210	

cct tgc aca gtg tac gac cgt ctg aaa gaa aca act gca ggc tac atc Pro Cys Thr Val Tyr Asp Arg Leu Lys Glu Thr Thr Ala Gly Tyr Ile 215 220 225 230	9128
act atg cac agg ccg aga ccg cac gct tat aca tcc tac ctg gaa gaa Thr Met His Arg Pro Arg Pro His Ala Tyr Thr Ser Tyr Leu Glu Glu 235 240 245	9176
tca tca ggg aaa gtt tac gca aag ccg cca tct ggg aag aac att acg Ser Ser Gly Lys Val Tyr Ala Lys Pro Pro Ser Gly Lys Asn Ile Thr 250 255 260	9224
tat gag tgc aag tgc ggc gac tac aag acc gga acc gtt tcg acc cgc Tyr Glu Cys Lys Cys Gly Asp Tyr Lys Thr Gly Thr Val Ser Thr Arg 265 270 275	9272
acc gaa atc act ggt tgc acc gcc atc aag cag tgc gtc gcc tat aag Thr Glu Ile Thr Gly Cys Thr Ala Ile Lys Gln Cys Val Ala Tyr Lys 280 285 290	9320
agc gac caa acg aag tgg gtc ttc aac tca ccg gac ttg atc aga cat Ser Asp Gln Thr Lys Trp Val Phe Asn Ser Pro Asp Leu Ile Arg His 295 300 305 310	9368
gac gac cac acg gcc caa ggg aaa ttg cat ttg cct ttc aag ttg atc Asp Asp His Thr Ala Gln Gly Lys Leu His Leu Pro Phe Lys Leu Ile 315 320 325	9416
ccg agt acc tgc atg gtc cct gtt gcc cac gcg ccg aat gta ata cat Pro Ser Thr Cys Met Val Pro Val Ala His Ala Pro Asn Val Ile His 330 335 340	9464
ggc ttt aaa cac atc agc ctc caa tta gat aca gac cac ttg aca ttg Gly Phe Lys His Ile Ser Leu Gln Leu Asp Thr Asp His Leu Thr Leu 345 350 355	9512
ctc acc acc agg aga cta ggg gca aac ccg gaa cca acc act gaa tgg Leu Thr Thr Arg Arg Leu Gly Ala Asn Pro Glu Pro Thr Thr Glu Trp 360 365 370	9560
atc gtc gga aag acg gtc aga aac ttc acc gtc gac cga gat ggc ctg Ile Val Gly Lys Thr Val Arg Asn Phe Thr Val Asp Arg Asp Gly Leu 375 380 385 390	9608
gaa tac ata tgg gga aat cat gag cca gtg agg gtc tat gcc caa gag Glu Tyr Ile Trp Gly Asn His Glu Pro Val Arg Val Tyr Ala Gln Glu 395 400 405	9656
tca gca cca gga gac cct cac gga tgg cca cac gaa ata gta cag cat Ser Ala Pro Gly Asp Pro His Gly Trp Pro His Glu Ile Val Gln His 410 415 420	9704
tac tac cat cgc cat cct gtg tac acc atc tta gcc gtc gca tca gct Tyr Tyr His Arg His Pro Val Tyr Thr Ile Leu Ala Val Ala Ser Ala 425 430 435	9752
acc gtg gcg atg atg att ggc gta act gtt gca gtg tta tgt gcc tgt	9800

Thr	Val	Ala	Met	Met	Ile	Gly	Val	Thr	Val	Ala	Val	Leu	Cys	Ala	Cys	
440						445					450					

aaa	gcg	cgc	cgt	gag	tgc	ctg	acg	cca	tac	gcc	ctg	gcc	cca	aac	gcc	9848
Lys	Ala	Arg	Arg	Glu	Cys	Leu	Thr	Pro	Tyr	Ala	Leu	Ala	Pro	Asn	Ala	
455					460					465					470	

gta	atc	cca	act	tcg	ctg	gca	ctc	ttg	tgc	tgc	gtt	agg	tcg	gcc	aat	9896
Val	Ile	Pro	Thr	Ser	Leu	Ala	Leu	Leu	Cys	Cys	Val	Arg	Ser	Ala	Asn	
			475						480					485		

gct gaaacgttca ccgagaccat gagttacttg tggtcgaaca gtcagccgtt															9949
Ala															

cttctggggtc	cagttgtgca	tacctttggc	cgctttcatc	gttctaattgc	gctgctgctc	10009
ctgctgcctg	ccttttttag	tggttgccgg	cgcctacctg	gcgaaggtag	acgcctacga	10069
acatgcgacc	actgttccaa	atgtgccaca	gataccgtat	aaggcacttg	ttgaaagggc	10129
agggtatgcc	ccgctcaatt	tggagatcac	tgtcatgtcc	tcggagggtt	tgccttccac	10189
caaccaagag	tacattacct	gcaaattcac	cactgtgggtc	ccctccccaa	aaatcaaattg	10249
ctgcggctcc	ttggaatgtc	agccggccgc	tcattgcagac	tataacctgca	aggtcttcgg	10309
aggggtctac	ccctttatgt	ggggaggagc	gcaatgtttt	tgcgacagtg	agaacagcca	10369
gatgagtgag	gcgtacgtcg	aattgtcagc	agattgcgcg	tctgaccacg	cgcaggcgat	10429
taaggtgcac	actgccgcga	tgaaagtagg	actgcgtatt	gtgtacggga	acactaccag	10489
tttcctagat	gtgtacgtga	acggagtcac	accaggaacg	tctaaagact	tgaaagtcat	10549
agctggacca	atttcagcat	cgtttacgcc	attcgatcat	aaggtcgtta	tccatcgcg	10609
cctggtgtac	aactatgact	tcccgggaata	tggagcgatg	aaaccaggag	cgtttggaga	10669
cattcaagct	acctccttga	ctagcaagga	tctcatcgcc	agcacagaca	ttaggctact	10729
caagccttcc	gccaagaacg	tgcattgtccc	gtacacgcag	gcctcatcag	gatttgagat	10789
gtggaaaaac	aactcaggcc	gccactgca	ggaaaccgca	cctttcgggt	gtaagattgc	10849
agtaaattccg	ctccgagcgg	tggactgttc	atacgggaac	attcccattt	ctattgacat	10909
cccgaacgct	gcctttatca	ggacatcaga	tgcaccactg	gtctcaacag	tcaaattgtga	10969
agtcagtgag	tgcacttatt	cagcagactt	cggcgggatg	gccaccctgc	agtatgtatc	11029
cgaccgcgaa	ggtcaatgcc	ccgtacattc	gcattcgagc	acagcaactc	tccaagagtc	11089
gacagtacat	gtcctggaga	aaggagcgg	gacagtacac	tttagcaccg	cgagtccaca	11149
ggcgaacttt	atcgtatcgc	tgtgtgggaa	gaagacaaca	tgcaatgcag	aatgtaaacc	11209

accagctgac catatcgtga gcaccccgca caaaaatgac caagaatttc aagccgccat 11269
ctcaaaaaca tcatggagtt ggctgtttgc ccttttcggc ggcgcctcgt cgctattaat 11329
tataggactt atgatttttg cttgcagcat gatgctgact agcacacgaa gatgaccgct 11389
acgcccgaat gatccgacca gcaaaactcg atgtacttcc gaggaactga tgtgcataat 11449
gcatcaggct ggtacattag atccccgctt accgcgggca atatagcaac actaaaaact 11509
cgatgtactt ccgaggaagc gcagtgcata atgctgcgca gtgttgccac ataaccacta 11569
tattaaccat ttatctagcg gacgccaaaa actcaatgta tttctgagga agcgtggtgc 11629
ataatgccac gcagcgtctg cataactttt attatttctt ttattaatca acaaaatttt 11689
gtttttaaca tttc 11703

<210> 16
<211> 64
<212> PRT
<213> Sindbis virus

<400> 16

Ser Ala Ala Pro Leu Val Thr Ala Met Cys Leu Leu Gly Asn Val Ser
1 5 10 15

Phe Pro Cys Asp Arg Pro Pro Thr Cys Tyr Thr Arg Glu Pro Ser Arg
20 25 30

Ala Leu Asp Ile Leu Glu Glu Asn Val Asn His Glu Ala Tyr Asp Thr
35 40 45

Leu Leu Asn Ala Ile Leu Arg Cys Gly Ser Ser Gly Arg Ser Lys Arg
50 55 60

<210> 17
<211> 423
<212> PRT
<213> Sindbis virus

<400> 17

Ser Val Ile Asp Asp Phe Thr Leu Thr Ser Pro Tyr Leu Gly Thr Cys
1 5 10 15

Ser Tyr Cys His His Thr Val Pro Cys Phe Ser Pro Val Lys Ile Glu
20 25 30

Gln	Val	Trp	Asp	Glu	Ala	Asp	Asp	Asn	Thr	Ile	Arg	Ile	Gln	Thr	Ser	35	40	45	
Ala	Gln	Phe	Gly	Tyr	Asp	Gln	Ser	Gly	Ala	Ala	Ser	Ala	Asn	Lys	Tyr	50	55	60	
Arg	Tyr	Met	Ser	Leu	Lys	Gln	Asp	His	Thr	Val	Lys	Glu	Gly	Thr	Met	65	70	75	80
Asp	Asp	Ile	Lys	Ile	Ser	Thr	Ser	Gly	Pro	Cys	Arg	Arg	Leu	Ser	Tyr	85	90	95	
Lys	Gly	Tyr	Phe	Leu	Leu	Ala	Lys	Cys	Pro	Pro	Gly	Asp	Ser	Val	Thr	100	105	110	
Val	Ser	Ile	Val	Ser	Ser	Asn	Ser	Ala	Thr	Ser	Cys	Thr	Leu	Ala	Arg	115	120	125	
Lys	Ile	Lys	Pro	Lys	Phe	Val	Gly	Arg	Glu	Lys	Tyr	Asp	Leu	Pro	Pro	130	135	140	
Val	His	Gly	Lys	Lys	Ile	Pro	Cys	Thr	Val	Tyr	Asp	Arg	Leu	Lys	Glu	145	150	155	160
Thr	Thr	Ala	Gly	Tyr	Ile	Thr	Met	His	Arg	Pro	Arg	Pro	His	Ala	Tyr	165	170	175	
Thr	Ser	Tyr	Leu	Glu	Glu	Ser	Ser	Gly	Lys	Val	Tyr	Ala	Lys	Pro	Pro	180	185	190	
Ser	Gly	Lys	Asn	Ile	Thr	Tyr	Glu	Cys	Lys	Cys	Gly	Asp	Tyr	Lys	Thr	195	200	205	
Gly	Thr	Val	Ser	Thr	Arg	Thr	Glu	Ile	Thr	Gly	Cys	Thr	Ala	Ile	Lys	210	215	220	
Gln	Cys	Val	Ala	Tyr	Lys	Ser	Asp	Gln	Thr	Lys	Trp	Val	Phe	Asn	Ser	225	230	235	240
Pro	Asp	Leu	Ile	Arg	His	Asp	Asp	His	Thr	Ala	Gln	Gly	Lys	Leu	His	245	250	255	
Leu	Pro	Phe	Lys	Leu	Ile	Pro	Ser	Thr	Cys	Met	Val	Pro	Val	Ala	His				

260					265					270					
Ala	Pro	Asn	Val	Ile	His	Gly	Phe	Lys	His	Ile	Ser	Leu	Gln	Leu	Asp
		275					280					285			
Thr	Asp	His	Leu	Thr	Leu	Leu	Thr	Thr	Arg	Arg	Leu	Gly	Ala	Asn	Pro
	290					295					300				
Glu	Pro	Thr	Thr	Glu	Trp	Ile	Val	Gly	Lys	Thr	Val	Arg	Asn	Phe	Thr
305						310					315				320
Val	Asp	Arg	Asp	Gly	Leu	Glu	Tyr	Ile	Trp	Gly	Asn	His	Glu	Pro	Val
				325					330					335	
Arg	Val	Tyr	Ala	Gln	Glu	Ser	Ala	Pro	Gly	Asp	Pro	His	Gly	Trp	Pro
			340					345					350		
His	Glu	Ile	Val	Gln	His	Tyr	Tyr	His	Arg	His	Pro	Val	Tyr	Thr	Ile
		355					360					365			
Leu	Ala	Val	Ala	Ser	Ala	Thr	Val	Ala	Met	Met	Ile	Gly	Val	Thr	Val
	370					375					380				
Ala	Val	Leu	Cys	Ala	Cys	Lys	Ala	Arg	Arg	Glu	Cys	Leu	Thr	Pro	Tyr
385						390					395				400
Ala	Leu	Ala	Pro	Asn	Ala	Val	Ile	Pro	Thr	Ser	Leu	Ala	Leu	Leu	Cys
				405					410					415	
Cys	Val	Arg	Ser	Ala	Asn	Ala									
				420											

<210> 18
 <211> 51
 <212> DNA
 <213> Artificial sequence

<220>
 <223> Linker

<220>
 <221> CDS
 <222> (1)..(51)

<400> 18

aga tct ggt ggc ggt ggc tcg ggc ggt ggt ggg tcg ggt ggc ggc gga 48
Arg Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly
1 5 10 15

tct	51
Ser	

```
<210> 19
<211> 17
<212> PRT
<213> Artificial sequence
```

<220>
<223> Synthetic Construct

<400> 19

Arg Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Gly Ser Gly Gly Gly Gly
1 5 10 15

Ser

<210>	20
<211>	45
<212>	DNA
<213>	Artificial sequence

```
<220>
<223> Linker
```

```
<220>
<221> CDS
<222> (1)..(45)
```

```
<400> 20
ggg ggc ggt ggc tcg ggc ggt ggt ggg tcg ggt ggc ggc gga tct
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
1 5 10 15
45
```

```
<210> 21
<211> 15
<212> PRT
<213> Artificial sequence
```

<220>
<223> Synthetic Construct

<400> 21

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

```

1           5           10           15

<210>  22
<211>  1050
<212>  DNA
<213>  Bluetongue virus 10

<220>
<221>  CDS
<222>  (1)..(1050)
<223>  VP7 gene

<400>  22
atg gac act atc gcc gca aga gca ctc act gtg atg cga gca tgt gct      48
Met Asp Thr Ile Ala Ala Arg Ala Leu Thr Val Met Arg Ala Cys Ala
1           5           10           15

acg ctt caa gag gca aga att gtg ttg gaa gcc aat gtg atg gaa att      96
Thr Leu Gln Glu Ala Arg Ile Val Leu Glu Ala Asn Val Met Glu Ile
           20           25           30

ttg ggg ata gct atc aat agg tac aat gga ctc act tta cga gga gtg      144
Leu Gly Ile Ala Ile Asn Arg Tyr Asn Gly Leu Thr Leu Arg Gly Val
           35           40           45

acg atg cgc ccg acc tcg tta gca caa aga aat gag atg ttt ttt atg      192
Thr Met Arg Pro Thr Ser Leu Ala Gln Arg Asn Glu Met Phe Phe Met
           50           55           60

tgt ttg gat atg atg ctg tct gct gct ggg ata aat gtt gga ccg ata      240
Cys Leu Asp Met Met Leu Ser Ala Ala Gly Ile Asn Val Gly Pro Ile
65           70           75           80

tcg cca gac tat act caa cat atg gct acg att ggt gta cta gca aca      288
Ser Pro Asp Tyr Thr Gln His Met Ala Thr Ile Gly Val Leu Ala Thr
           85           90           95

ccg gaa ata cct ttt aca acg gaa gcg gcg aat gaa ata gca cga gtg      336
Pro Glu Ile Pro Phe Thr Thr Glu Ala Ala Asn Glu Ile Ala Arg Val
           100          105          110

act ggg gag act tcg aca tgg ggg cca gcg cgt cag cct tat ggt ttc      384
Thr Gly Glu Thr Ser Thr Trp Gly Pro Ala Arg Gln Pro Tyr Gly Phe
           115          120          125

ttc ctt gaa act gag gaa acc ttc caa cca ggg agg tgg ttc atg cgc      432
Phe Leu Glu Thr Glu Glu Thr Phe Gln Pro Gly Arg Trp Phe Met Arg
           130          135          140

gcc gct caa gca gta act gca gta gtg tgc ggt ccg gat atg att caa      480
Ala Ala Gln Ala Val Thr Ala Val Val Cys Gly Pro Asp Met Ile Gln
145          150          155          160

gtg tca ctt aat gct gga gcg aga gga gat gta caa cag ata ttt cag      528
Val Ser Leu Asn Ala Gly Ala Arg Gly Asp Val Gln Gln Ile Phe Gln

```

165										170					175					
ggt	cgt	aat	gat	ccc	atg	atg	ata	tat	tta	gtg	tgg	agg	aga	atc	gaa	576				
Gly	Arg	Asn	Asp	Pro	Met	Met	Ile	Tyr	Leu	Val	Trp	Arg	Arg	Ile	Glu					
			180					185						190						
aac	ttt	gcg	atg	gcg	caa	ggt	aat	tca	cag	caa	act	caa	gcg	ggt	gtg	624				
Asn	Phe	Ala	Met	Ala	Gln	Gly	Asn	Ser	Gln	Gln	Thr	Gln	Ala	Gly	Val					
		195					200					205								
act	gtc	agt	gtt	ggt	gga	gtt	gac	atg	agg	gcg	gga	cgc	att	ata	gcg	672				
Thr	Val	Ser	Val	Gly	Gly	Val	Asp	Met	Arg	Ala	Gly	Arg	Ile	Ile	Ala					
	210					215					220									
tgg	gat	gga	cag	gcc	gcg	ctg	cat	gtg	cat	aat	ccg	aca	caa	cag	aat	720				
Trp	Asp	Gly	Gln	Ala	Ala	Leu	His	Val	His	Asn	Pro	Thr	Gln	Gln	Asn					
225					230					235					240					
gcg	atg	gtg	caa	ata	cag	gtt	gtg	ttc	tat	ata	tct	atg	gat	aaa	act	768				
Ala	Met	Val	Gln	Ile	Gln	Val	Val	Phe	Tyr	Ile	Ser	Met	Asp	Lys	Thr					
				245					250					255						
tta	aac	cag	tac	ccc	gct	ttg	act	gct	gag	att	ttc	aat	gtt	tac	agc	816				
Leu	Asn	Gln	Tyr	Pro	Ala	Leu	Thr	Ala	Glu	Ile	Phe	Asn	Val	Tyr	Ser					
			260					265					270							
ttc	agg	gac	cac	aca	tgg	cat	ggg	cta	aga	acg	gcg	ata	tta	aac	aga	864				
Phe	Arg	Asp	His	Thr	Trp	His	Gly	Leu	Arg	Thr	Ala	Ile	Leu	Asn	Arg					
		275					280					285								
acc	aca	ctg	cca	aac	atg	ctg	cca	cca	atc	ttc	cca	cca	aat	gat	cga	912				
Thr	Thr	Leu	Pro	Asn	Met	Leu	Pro	Pro	Ile	Phe	Pro	Pro	Asn	Asp	Arg					
		290				295					300									
gat	agc	atc	tta	act	ctt	cta	ctt	tta	tct	aca	ctt	gct	gat	gtt	tac	960				
Asp	Ser	Ile	Leu	Thr	Leu	Leu	Leu	Leu	Ser	Thr	Leu	Ala	Asp	Val	Tyr					
305					310					315					320					
act	gtt	tta	agg	cca	gag	ttt	gcg	att	cac	ggc	gta	aat	ccg	atg	cca	1008				
Thr	Val	Leu	Arg	Pro	Glu	Phe	Ala	Ile	His	Gly	Val	Asn	Pro	Met	Pro					
				325					330					335						
ggg	ccg	ctc	aca	cgt	gct	att	gcg	cgc	gcc	gcc	tat	gtg	tag			1050				
Gly	Pro	Leu	Thr	Arg	Ala	Ile	Ala	Arg	Ala	Ala	Tyr	Val								
			340					345												

<210> 23
 <211> 349
 <212> PRT
 <213> Bluetongue virus 10

<400> 23

Met	Asp	Thr	Ile	Ala	Ala	Arg	Ala	Leu	Thr	Val	Met	Arg	Ala	Cys	Ala
1				5					10					15	

Thr Leu Gln Glu Ala Arg Ile Val Leu Glu Ala Asn Val Met Glu Ile
 20 25 30

Leu Gly Ile Ala Ile Asn Arg Tyr Asn Gly Leu Thr Leu Arg Gly Val
 35 40 45

Thr Met Arg Pro Thr Ser Leu Ala Gln Arg Asn Glu Met Phe Phe Met
 50 55 60

Cys Leu Asp Met Met Leu Ser Ala Ala Gly Ile Asn Val Gly Pro Ile
 65 70 75 80

Ser Pro Asp Tyr Thr Gln His Met Ala Thr Ile Gly Val Leu Ala Thr
 85 90 95

Pro Glu Ile Pro Phe Thr Thr Glu Ala Ala Asn Glu Ile Ala Arg Val
 100 105 110

Thr Gly Glu Thr Ser Thr Trp Gly Pro Ala Arg Gln Pro Tyr Gly Phe
 115 120 125

Phe Leu Glu Thr Glu Glu Thr Phe Gln Pro Gly Arg Trp Phe Met Arg
 130 135 140

Ala Ala Gln Ala Val Thr Ala Val Val Cys Gly Pro Asp Met Ile Gln
 145 150 155 160

Val Ser Leu Asn Ala Gly Ala Arg Gly Asp Val Gln Gln Ile Phe Gln
 165 170 175

Gly Arg Asn Asp Pro Met Met Ile Tyr Leu Val Trp Arg Arg Ile Glu
 180 185 190

Asn Phe Ala Met Ala Gln Gly Asn Ser Gln Gln Thr Gln Ala Gly Val
 195 200 205

Thr Val Ser Val Gly Gly Val Asp Met Arg Ala Gly Arg Ile Ile Ala
 210 215 220

Trp Asp Gly Gln Ala Ala Leu His Val His Asn Pro Thr Gln Gln Asn
 225 230 235 240

Ala Met Val Gln Ile Gln Val Val Phe Tyr Ile Ser Met Asp Lys Thr
245 250 255

Leu Asn Gln Tyr Pro Ala Leu Thr Ala Glu Ile Phe Asn Val Tyr Ser
260 265 270

Phe Arg Asp His Thr Trp His Gly Leu Arg Thr Ala Ile Leu Asn Arg
275 280 285

Thr Thr Leu Pro Asn Met Leu Pro Pro Ile Phe Pro Pro Asn Asp Arg
290 295 300

Asp Ser Ile Leu Thr Leu Leu Leu Leu Ser Thr Leu Ala Asp Val Tyr
305 310 315 320

Thr Val Leu Arg Pro Glu Phe Ala Ile His Gly Val Asn Pro Met Pro
325 330 335

Gly Pro Leu Thr Arg Ala Ile Ala Arg Ala Ala Tyr Val
340 345

<210> 24
<211> 585
<212> DNA
<213> *Coccidioides immitis*

<220>
<221> CDS
<222> (1)..(585)
<223> Ag2/PRA gene

<400> 24
atg cag ttc tct cac gct ctc atc gct ctc gtc gct gcc ggc ctc gcc 48
Met Gln Phe Ser His Ala Leu Ile Ala Leu Val Ala Ala Gly Leu Ala
1 5 10 15
agt gcc cag ctc cca gac atc cca cct tgc gct ctc aac tgc ttc gtt 96
Ser Ala Gln Leu Pro Asp Ile Pro Pro Cys Ala Leu Asn Cys Phe Val
20 25 30
gag gct ctc ggc aac gat ggc tgc act cgc ttg acc gac ttc aag tgc 144
Glu Ala Leu Gly Asn Asp Gly Cys Thr Arg Leu Thr Asp Phe Lys Cys
35 40 45
cac tgc tcc aag cct gag cta cca gga cag atc act cct tgc gtt gag 192
His Cys Ser Lys Pro Glu Leu Pro Gly Gln Ile Thr Pro Cys Val Glu
50 55 60
gag gcc tgc cct ctc gac gcc cgt atc tcc gtc tcc aac atc gtc gtt 240

Glu	Ala	Cys	Pro	Leu	Asp	Ala	Arg	Ile	Ser	Val	Ser	Asn	Ile	Val	Val		
65					70					75					80		
gac	cag	tgc	tcc	aag	gcc	ggg	gtc	cca	att	gac	atc	cca	cca	gtt	gac		288
Asp	Gln	Cys	Ser	Lys	Ala	Gly	Val	Pro	Ile	Asp	Ile	Pro	Pro	Val	Asp		
				85					90					95			
acc	acc	gcc	gct	ccc	gag	cca	tcc	gag	acc	gct	gag	ccc	acc	gct	gag		336
Thr	Thr	Ala	Ala	Pro	Glu	Pro	Ser	Glu	Thr	Ala	Glu	Pro	Thr	Ala	Glu		
				100				105					110				
cca	acc	gag	gag	ccc	act	gcc	gag	cct	acc	gct	gag	ccc	acc	gct	gag		384
Pro	Thr	Glu	Glu	Pro	Thr	Ala	Glu	Pro	Thr	Ala	Glu	Pro	Thr	Ala	Glu		
				115			120					125					
ccg	act	cat	gag	ccc	acc	gag	gag	ccc	act	gcc	gtc	cca	acc	ggc	act		432
Pro	Thr	His	Glu	Pro	Thr	Glu	Glu	Pro	Thr	Ala	Val	Pro	Thr	Gly	Thr		
				130			135				140						
ggc	ggg	ggg	gtc	ccc	act	ggc	acc	ggg	tcc	ttc	acc	gtc	act	ggc	aga		480
Gly	Gly	Gly	Val	Pro	Thr	Gly	Thr	Gly	Ser	Phe	Thr	Val	Thr	Gly	Arg		
145					150				155					160			
cca	act	gcc	tcc	acc	cca	gct	gag	ttc	cca	ggg	gct	ggc	tcc	aac	gtc		528
Pro	Thr	Ala	Ser	Thr	Pro	Ala	Glu	Phe	Pro	Gly	Ala	Gly	Ser	Asn	Val		
				165				170						175			
cgt	gcc	agc	gtt	ggc	ggc	att	gct	gct	gct	ctc	ctc	ggg	ctc	gct	gcc		576
Arg	Ala	Ser	Val	Gly	Gly	Ile	Ala	Ala	Ala	Leu	Leu	Gly	Leu	Ala	Ala		
			180				185						190				
tac	ctg	taa															585
Tyr	Leu																

<210> 25
 <211> 194
 <212> PRT
 <213> Coccidioides immitis

<400> 25

Met	Gln	Phe	Ser	His	Ala	Leu	Ile	Ala	Leu	Val	Ala	Ala	Gly	Leu	Ala		
1				5					10					15			

Ser	Ala	Gln	Leu	Pro	Asp	Ile	Pro	Pro	Cys	Ala	Leu	Asn	Cys	Phe	Val		
			20				25						30				

Glu	Ala	Leu	Gly	Asn	Asp	Gly	Cys	Thr	Arg	Leu	Thr	Asp	Phe	Lys	Cys		
		35				40						45					

His	Cys	Ser	Lys	Pro	Glu	Leu	Pro	Gly	Gln	Ile	Thr	Pro	Cys	Val	Glu		
	50					55					60						

Glu Ala Cys Pro Leu Asp Ala Arg Ile Ser Val Ser Asn Ile Val Val
 65 70 75 80

Asp Gln Cys Ser Lys Ala Gly Val Pro Ile Asp Ile Pro Pro Val Asp
 85 90 95

Thr Thr Ala Ala Pro Glu Pro Ser Glu Thr Ala Glu Pro Thr Ala Glu
 100 105 110

Pro Thr Glu Glu Pro Thr Ala Glu Pro Thr Ala Glu Pro Thr Ala Glu
 115 120 125

Pro Thr His Glu Pro Thr Glu Glu Pro Thr Ala Val Pro Thr Gly Thr
 130 135 140

Gly Gly Gly Val Pro Thr Gly Thr Gly Ser Phe Thr Val Thr Gly Arg
 145 150 155 160

Pro Thr Ala Ser Thr Pro Ala Glu Phe Pro Gly Ala Gly Ser Asn Val
 165 170 175

Arg Ala Ser Val Gly Gly Ile Ala Ala Ala Leu Leu Gly Leu Ala Ala
 180 185 190

Tyr Leu

<210> 26
 <211> 906
 <212> DNA
 <213> Streptococcus pneumoniae

<220>
 <221> CDS
 <222> (1)..(906)
 <223> PspA gene

<400> 26
 gaa gaa tct ccc gta gcc agt cag tct aaa gct gag aaa gac tat gat 48
 Glu Glu Ser Pro Val Ala Ser Gln Ser Lys Ala Glu Lys Asp Tyr Asp
 1 5 10 15

gca gcg aag aaa gat gct aag aat gcg aaa aaa gca gta gaa gat gct 96
 Ala Ala Lys Lys Asp Ala Lys Asn Ala Lys Lys Ala Val Glu Asp Ala
 20 25 30

caa aag gct tta gat gat gca aaa gct gct cag aaa aaa tat gac gag Gln Lys Ala Leu Asp Asp Ala Lys Ala Ala Gln Lys Lys Tyr Asp Glu 35 40 45	144
gat cag aag aaa act gag gag aaa gcc gcg cta gaa aaa gca gcg tct Asp Gln Lys Lys Thr Glu Glu Lys Ala Ala Leu Glu Lys Ala Ala Ser 50 55 60	192
gaa gag atg gat aag gca gtg gca gca gtt caa caa gcg tat cta gcc Glu Glu Met Asp Lys Ala Val Ala Ala Val Gln Gln Ala Tyr Leu Ala 65 70 75 80	240
tat caa caa gct aca gac aaa gcc gca aaa gac gca gca gat aag atg Tyr Gln Gln Ala Thr Asp Lys Ala Ala Lys Asp Ala Ala Asp Lys Met 85 90 95	288
ata gat gaa gct aag aaa cgc gaa gaa gag gca aaa act aaa ttt aat Ile Asp Glu Ala Lys Lys Arg Glu Glu Glu Ala Lys Thr Lys Phe Asn 100 105 110	336
act gtt cga gca atg gta gtt cct gag cca gag cag ttg gct gag act Thr Val Arg Ala Met Val Val Pro Glu Pro Glu Gln Leu Ala Glu Thr 115 120 125	384
aag aaa aaa tca gaa gaa gct aaa caa aaa gca cca gaa ctt act aaa Lys Lys Lys Ser Glu Glu Ala Lys Gln Lys Ala Pro Glu Leu Thr Lys 130 135 140	432
aaa cta gaa gaa gct aaa gca aaa tta gaa gag gct gag aaa aaa gct Lys Leu Glu Glu Ala Lys Ala Lys Leu Glu Glu Ala Glu Lys Lys Ala 145 150 155 160	480
act gaa gcc aaa caa aaa gtg gat gct gaa gaa gtc gct cct caa gct Thr Glu Ala Lys Gln Lys Val Asp Ala Glu Glu Val Ala Pro Gln Ala 165 170 175	528
aaa atc gct gaa ttg gaa aat caa gtt cat aga cta gaa caa gag ctc Lys Ile Ala Glu Leu Glu Asn Gln Val His Arg Leu Glu Gln Glu Leu 180 185 190	576
aaa gag att gat gag tct gaa tca gaa gat tat gct aaa gaa ggt ttc Lys Glu Ile Asp Glu Ser Glu Ser Glu Asp Tyr Ala Lys Glu Gly Phe 195 200 205	624
cgt gct cct ctt caa tct aaa ttg gat gcc aaa aaa gct aaa cta tca Arg Ala Pro Leu Gln Ser Lys Leu Asp Ala Lys Lys Ala Lys Leu Ser 210 215 220	672
aaa ctt gaa gag tta agt gat aag att gat gag tta gac gct gaa att Lys Leu Glu Glu Leu Ser Asp Lys Ile Asp Glu Leu Asp Ala Glu Ile 225 230 235 240	720
gca aaa ctt gaa gat caa ctt aaa gct gct gaa gaa aac aat aat gta Ala Lys Leu Glu Asp Gln Leu Lys Ala Ala Glu Glu Asn Asn Asn Val 245 250 255	768

gaa gac tac ttt aaa gaa ggt tta gag aaa act att gct gct aaa aaa 816
 Glu Asp Tyr Phe Lys Glu Gly Leu Glu Lys Thr Ile Ala Ala Lys Lys
 260 265 270

gct gaa tta gaa aaa act gaa gct gac ctt aag aaa gca gtt aat gag 864
 Ala Glu Leu Glu Lys Thr Glu Ala Asp Leu Lys Lys Ala Val Asn Glu
 275 280 285

cca gaa aaa cca gct cca gct cca gaa act cca gcc cca gaa 906
 Pro Glu Lys Pro Ala Pro Ala Pro Glu Thr Pro Ala Pro Glu
 290 295 300

<210> 27
 <211> 302
 <212> PRT
 <213> Streptococcus pneumoniae

 <400> 27

Glu Glu Ser Pro Val Ala Ser Gln Ser Lys Ala Glu Lys Asp Tyr Asp
 1 5 10 15

Ala Ala Lys Lys Asp Ala Lys Asn Ala Lys Lys Ala Val Glu Asp Ala
 20 25 30

Gln Lys Ala Leu Asp Asp Ala Lys Ala Ala Gln Lys Lys Tyr Asp Glu
 35 40 45

Asp Gln Lys Lys Thr Glu Glu Lys Ala Ala Leu Glu Lys Ala Ala Ser
 50 55 60

Glu Glu Met Asp Lys Ala Val Ala Ala Val Gln Gln Ala Tyr Leu Ala
 65 70 75 80

Tyr Gln Gln Ala Thr Asp Lys Ala Ala Lys Asp Ala Ala Asp Lys Met
 85 90 95

Ile Asp Glu Ala Lys Lys Arg Glu Glu Glu Ala Lys Thr Lys Phe Asn
 100 105 110

Thr Val Arg Ala Met Val Val Pro Glu Pro Glu Gln Leu Ala Glu Thr
 115 120 125

Lys Lys Lys Ser Glu Glu Ala Lys Gln Lys Ala Pro Glu Leu Thr Lys
 130 135 140

Lys Leu Glu Glu Ala Lys Ala Lys Leu Glu Glu Ala Glu Lys Lys Ala

145		150		155		160									
Thr	Glu	Ala	Lys	Gln	Lys	Val	Asp	Ala	Glu	Glu	Val	Ala	Pro	Gln	Ala
				165					170					175	
Lys	Ile	Ala	Glu	Leu	Glu	Asn	Gln	Val	His	Arg	Leu	Glu	Gln	Glu	Leu
			180					185					190		
Lys	Glu	Ile	Asp	Glu	Ser	Glu	Ser	Glu	Asp	Tyr	Ala	Lys	Glu	Gly	Phe
		195					200					205			
Arg	Ala	Pro	Leu	Gln	Ser	Lys	Leu	Asp	Ala	Lys	Lys	Ala	Lys	Leu	Ser
	210					215					220				
Lys	Leu	Glu	Glu	Leu	Ser	Asp	Lys	Ile	Asp	Glu	Leu	Asp	Ala	Glu	Ile
225					230					235					240
Ala	Lys	Leu	Glu	Asp	Gln	Leu	Lys	Ala	Ala	Glu	Glu	Asn	Asn	Asn	Val
				245					250					255	
Glu	Asp	Tyr	Phe	Lys	Glu	Gly	Leu	Glu	Lys	Thr	Ile	Ala	Ala	Lys	Lys
			260					265					270		
Ala	Glu	Leu	Glu	Lys	Thr	Glu	Ala	Asp	Leu	Lys	Lys	Ala	Val	Asn	Glu
		275					280					285			
Pro	Glu	Lys	Pro	Ala	Pro	Ala	Pro	Glu	Thr	Pro	Ala	Pro	Glu		
	290					295					300				